Convergence to a Convention in 2x2 Games with Adaptive Play

Ethan Holdahl

University of Oregon March 7, 2023

1 Introduction

[NEEDS WORK AND LIT REVIEW.... IGNORE FOR NOW]

With over 1000 citations, Young [1993] is a seminal paper in the field of evolutionary game theory. In it, Young introduces a model of learning called adaptive play in which players best respond to the sampled history of play. Though adaptive play, Young establishes that under certain conditions, a convention, a self enforcing pattern of play, will emerge. Theorem 1 proved that in weakly acyclic n-person games, if the information gathered in the available history by the players is sufficiently incomplete, that play will eventually converge to a convention. The rest of the paper builds upon this theorem and defines which conventions are stochastically stable. In Young [1998], Young weakened the incomplete sampling requirement in coordination games to being less than 1/2 of the available history. Here, I show that in games where players have no more than 2 strategies, any degree of incomplete information is sufficient for the results in Young [1993]. Furthermore, I show that in all but certain formulations of the 2x2 game, no condition on the degree of incomplete information is necessary for conventions to emerge in adaptive play.

2 Proofs

Consider an infinitely repeated N player game where each player i plays one pure strategy from their set of strategies, $s_i \in S_i$ each period. The strategy-tuple $s(t) = (s_1(t), ..., s_N(t))$ is recorded and will be referred to as the play at time t. The history of plays up to and including time t is the ordered vector $h(t) = (s(1), s(2), s(3), ..., s(t))$. Fix integers k and m such that $1 \leq k \leq m$. In period $t+1$ when $t \geq m$, each player independently samples, without replacement, k of the m most recent plays from $h(t)$. Denote player i's sample in period $t+1$ as R_i^{t+1}. After drawing their samples, each player plays a best response to their sampled distribution of the other players' strategies, $s_i(t+1) \in \{BR_i(R_i^{t+1})\}$. A player's best response function selects the set of strategies that maximize their payoffs, $BR_i(R_i^{t+1}) = \max\limits_{\{s_i\}} \sum\limits_{s_{-i} \in S_{-i}} \pi_i(s_i, s_{-i}) p(s_{-i}|R_i^{t+1})$ where $\pi_i(s_i, s_{-i})$ is the payoff player i gets when they play s_i and the other players collectively play s_{-i}, and $p(s_{-i}|R_i^{t+1}) = \prod\limits_{s_j \in s_{-i}} p(s_j|R_i^{t+1})$ is the probability that each player other that player i plays their strategy s_j in s_{-i}, assuming players choose their strategies independently of one another.

The decision making process described above is called unperturbed adaptive play with memory size m and sample size k. Though an adaptive play process, self enforcing patterns of play, called conventions, can emerge.

Definition 1. *Convention: Any state h consisting of m repetitions of the same* *strict, pure strategy Nash equilibrium.*

In Young [1998], Young proved that in coordination games so long as $k \leq \frac{m}{2}$, a convention will eventually be reached in unperturbed adaptive play. I will introduce a Lemma below that will be used to show that in 2x2 coordination games, any degree of incomplete information is sufficient for a convention to eventually be reached.

Lemma 1 will be used in the proof of Theorems 1 and 2.

Lemma 1. *Let G be a 2x2 coordination game with unperturbed adaptive play with memory size m and sample size k. If in any period $t > m$ a strategy-tuple $s^* = (s_1^*, s_2^*)$ that constitutes a strict Nash equilibrium can be played as a best response to the sampled history, then there exists a positive probability that a convention of playing s can be reached.*

Proof of Lemma 1. Assume that at period t there exists a strategy-tuple $s^* = (s_1^*, s_2^*)$ that constitutes a strict Nash Equilibrium can be played as a best response to the sampled history. I will use proof by induction to show that there exists a positive probability that s^* is played in periods t through $t + m - 1$, which would result in m repetitions of a strict pure strategy Nash equilibrium in the history of play which is the definition of a convention.

Since

Base Step: By assumption, $s^* = (s_1^*, s_2^*)$ constitutes a Nash Equilibrium and can be played as a best response to the sampled history in period t.

Inductive Step: I show that $\forall e \geq t$ if s_1^* and s_2^* were played as a best responses for players 1 and 2 in period e then there exists a positive probability that s^* is played in period $e + 1$.

Assume s_1^* and s_2^* were played for players 1 and 2 respectively in period e. I must show that there exists a positive probability that s_1^* and s_2^* are played in period $e + 1$.

In period $e + 1$, it is possible that both players sample all but one of the same records of play that they did in the previous period, with the addition of the record of play $s(e) = s^* = (s_1^*, s_2^*)$. Let r_i^e denote the only record of play sampled by player i in period e but not period $e + 1$. Given that s_i^* was a best response to R_i^e, it suffices to show that if by replacing r_i^e with $s(e)$, the difference between player i's expected payoff for playing s_i^* and s_i', player i's other pure strategy, is non-decreasing then s_i^* is a best response to R_i^{e+1}.

By replacing r_i^e with $s(e)$, either the sampled distribution of the other player's strategies remain unchanged or the frequency of the strategy corresponding to the Nash equilibrium (s_1^*, s_2^*) increases. If the distribution remains unchanged then clearly the difference in expected payoffs is non-decreasing.

Since G is a 2x2 coordination game, that means that (s_1^*, s_2^*) and (s_1', s_2')

are both Nash equilibria. By definition this requires $\pi_1(s_1^*, s_2^*) \geq \pi_1(s_1', s_2^*)$ and $\pi_1(s_1^*, s_2') \leq \pi_1(s_1', s_2')$. Combining these inequalities $\pi_1(s_1^*, s_2^*) - \pi_1(s_1', s_2^*) \geq 0 \geq \pi_1(s_1^*, s_2') - \pi_1(s_1', s_2')$ so $\pi_1(s_1^*, s_2^*) - \pi_1(s_1^*, s_2') \geq \pi_1(s_1', s_2^*) - \pi_1(s_1', s_2')$. As such, if frequency of s_2^* increases in player 1's sample, and thus the frequency of s_2' decreases, then the difference in expected payoffs between s_1^* and s_1' for player 1 is also non-decreasing.

The same logic can be applied to player 2. As such, I have shown that if both players sample all but one of the same records of play that they did in the previous period, with the addition of the record of play $s(e) = s^* = (s_1^*, s_2^*)$ then s_i^* is a best response to R_i^{e+1}. Since both s_1^* and s_2^* are best replies to each respective player's sample in period $e + 1$, there is a positive probability that $s^* = (s_1^*, s_2^*)$ is played in period $e + 1$ as desired.

Since both the base step and the inductive step have been shown, by mathematical induction I have proven that if at any period $t > m$ there exists a strategy-tuple $s^* = (s_1^*, s_2^*)$ that constitutes a strict Nash Equilibrium can be played as a best response to the sampled history, then there exists a positive probability that a convention of playing s^* can be reached. $\square$

Remark of Lemma 1. Lemma 1 states that in 2x2 coordination games with unperturbed adaptive play with memory size m and sample size k. If in any period $t > m$ a strategy-tuple $s^* = (s_1^*, s_2^*)$ that constitutes a Nash equilibrium can be played as a best response to the sampled history, then there exists a positive probability that a convention of playing s^* can be reached. This is fairly intuitive in the 2x2 case, but the same logic does not necessarily extend to larger games. Consider, for example, the following 3x3 coordination game, Game 1:

Game 1

Player 2

		b_1	b_2	b_3
	a_1	$1, 1^*$	$0, 0$	$0, 0$
Player 1	a_2	$0, 0$	$1, 1^*$	$0, 0$
	a_3	$0, 0$	$-5, -5$	$1, 1^*$

Let $t > m$, $k = 7$, $R_1^t = (b_1, b_1, b_2, b_3, b_3, b_3, b_3)$ and $R_2^t = (b_1, b_1, b_2, b_2, b_2, b_2, b_3)$. Given the distribution of sampled play, player 1 has an expected payoff of 2/7 if they play a_1, 1/7 if they play a_2 and $-1/7$ if they play a_3. So they play a_1. Player 2 has an expected payoff of 2/7 if they play b_1, $-1/7$ if they play b_2 and 1/7 if they play b_3. So they play b_1. So $s^* = (a_1, b_1)$, which constitutes a Nash equilibrium, can be played as a best response to the sampled history. However, by replacing any record sampled in R_i^t with $s(t)$, a_1 and b_1 are not guaranteed to still be best responses. Specifically, if the record b_2 is removed from player 1's sample and replaced with b_1 then player 1's expected payoffs become 3/7 if they play a_1, 0 if they play a_2 and $4/7$ if they play a_3. So by replacing b_2 with b_1 in their sample, which previously yielded a_1 as the best response, now gives b_3 as the best response even though a_1 is the best response to b_1.

Therefore, the mechanisms that allow Lemma 1 to work do not necessarily apply to larger games.

Theorem 1. *Let G be a 2x2 coordination game with unperturbed adaptive play with memory size m and sample size k. If $k < m$ then from any initial state, the unperturbed process converges with probability one to a convention and locks in.*

Proof of Theorem 1. Assume both players only sample the most recent k records. If at any period, t, a strategy-tuple $s^* = (s_1^*, s_2^*)$ that constitutes a Nash Equilibrium can be played as a best response to the sampled play, then by Lemma 1, there exists a positive probability that a convention of playing s^* can be reached.

So, it remains to be proved that a convention can be reached if a strategy-tuple $s^* = (s_1^*, s_2^*)$ that constitutes a Nash Equilibrium could never be reached in a period if both players always sample the most recent k records.

Assume that if players sampled the most recent k records that they would never play a pair of best responses that constitute a Nash equilibrium. Since this is adaptive play in a 2x2 coordination game, we know that if players are not playing a Nash equilibrium that eventually one or both of their best responses will change to the coordinate with the strategy that the other player was playing. In order for players to never share a pair of best responses that constitute a Nash equilibrium, it must be the case that whenever either players' best response changes, the best response of the other player simultaneously changes as well. Let t be a period where both players would simultaneously switch their best response if they both sampled the most recent k records. In this period, there is a positive probability that player 1 samples the most recent k records and, since $m > k$, there is a positive probability that player 2 samples the most recent $k+1$ records excluding the most recent record. Consequently, player 1's best response changed to coordinate with player 2's play in the period $t-1$, but because player 2 sampled the same set of records that they did in period $t-1$, their best response did not change in period t. Consequently, in period t player 1 and player 2 best responses are some s_1^* and s_2^* respectively which constitute a Nash equilibrium. So, by Lemma 1, there exists a positive probability that a convention of playing $s^* = (s_1^*, s_2^*)$ can be reached.

I have proven that from any initial state when $k < m$ a convention can be reached with positive probability. Since a convention can be reached from any arbitrary state and since conventions are self enforcing, we know that as $T \to \infty$ that the most recent m records in $h(T)$, the state at time T, converges with probability one to a convention and locks in. ■

Theorem 2. *Let G be a 2x2 coordination game with unperturbed adaptive play with memory size m and sample size k. Let $s_1 = \{a_1, a_2\}$ and $s_2 = \{b_1, b_2\}$ be the set of strategies available to player 1 and 2 respectively such that (a_1, b_1) and (a_2, b_2) are both strict Nash equilibria. Let $\alpha \in (0,1)$ be the smallest probability that player 2 plays strategy b_1 such that player 1's best response is playing action a_1. Likewise, Let $\beta \in (0,1)$ be the smallest probability that player 1 plays action a_1 such that player 2's best response is playing action b_1.*

If $\lceil \alpha k \rceil \neq \lceil (1-\beta)k \rceil$ then from any initial state, the unperturbed adaptive play process converges with probability one to a convention and locks in.

Proof of Theorem 2. Due to Lemma 1, it is sufficient to show that if there exist a period t where the best response of both players constitute a Nash equilibrium then a convention can be reached.

I will show that when $\lceil \alpha k \rceil \neq \lceil (1-\beta)k \rceil$, that by sampling the most recent k records there will always exist a period t where the best response of both players constitute a Nash equilibrium.

I use proof by contradiction. Assume that when sampling the most recent k records, the best responses of both players never constitute a Nash equilibrium and that $\lceil \alpha k \rceil \neq \lceil (1-\beta)k \rceil$. Because this is unperturbed adaptive play in a coordination game, we know that if players aren't playing a Nash equilibrium that eventually one or both of their best responses will change to the strategy that the other player was playing. In order for the players' best responses to never constitute a Nash equilibrium, it must be the case that whenever either players' best response changes, the best response of the other player simultaneously changes as well. Given enough periods, this pattern of simultaneously switching strategies will yield a history of play of perfect miscoordination. This means that the amount of records of b_1 in player 1's sample necessary for player 1 to play a_1 as a best response is equal to the amount of records of a_2 in player 2's sample necessary for player 2 to play b_2 as a best response: $\lceil \alpha k \rceil = \lceil (1-\beta)k \rceil$. However, this contradicts the assumption that $\lceil \alpha k \rceil \neq \lceil (1-\beta)k \rceil$. So, when sampling the most recent k records, it is impossible for the best responses of both players to never constitute a Nash equilibrium if $\lceil \alpha k \rceil \neq \lceil (1-\beta)k \rceil$.

So, I have shown that when $\lceil \alpha k \rceil \neq \lceil (1-\beta)k \rceil$, if players only sampling the most recent k records there must exist a period t where the best response of both players constitute a Nash equilibrium. At which point Lemma 1 applies which means given any $k \leq m$ and starting from any given history, there is a positive probability that a convention can be reached. Since conventions are self enforcing, we know that as $T \to \infty$ that the most recent m records in $h(T)$, the state at time T, converges with probability one to a convention and locks in. ■

References

H. P. Young. The evolution of conventions. *Econometrica: Journal of the Econometric Society*, pages 57–84, 1993.

H. P. Young. *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton University Press, 1998. ISBN 9780691086873. URL http://www.jstor.org/stable/j.ctv10h9d35.